

when executed by an information handling system, causes the information handling system to mine threaded online discussions by performing actions comprising:

performing, by the information handling system, a natural language processing (NLP) analysis of one or more threaded discussions pertaining to a given topic, wherein the analysis is performed across one or more web sites with each of the web sites including one or more of the threaded discussions, wherein the analysis results in a plurality of harvested discussions;

identifying a question from the harvested discussions;

identifying a plurality of candidate answers from the harvested discussions, wherein each of the plurality of candidate answers pertain to the identified question;

aggregating and merging a selected plurality of harvested discussions corresponding to each of the candidate answers, wherein the selected plurality of harvested discussions are supporting evidence corresponding to the respective candidate answer;

generating a supporting evidence score based on one or more factors of the supporting evidence for each of the candidate answers; and

scoring each of the plurality of candidate answers, wherein the scoring calculates an overall score corresponding to each of the candidate answers, wherein the overall score is based upon at least the supporting evidence score.

15. The computer program product of claim **14** wherein the actions further comprise:

comparing a plurality of questions found in the threaded discussions to a posed question, wherein the identified question matches the posed question; and

adding one or more of the correlated harvested discussions to a corpus that is utilized in a deep question answering system.

16. The computer program product of claim **14** wherein the overall score is based upon one or more component scores selected from the group consisting of the supporting evidence score, an answer post score, a post provider score, and a follow-up score, and wherein the selected answer has the highest overall score when compared to the other candidate answers.

17. The computer program product of claim **16** wherein the actions further comprise:

performing at least one sentiment analysis on each of the candidate scores, wherein the sentiment analyses result in a sentiment analysis score corresponding to each of the candidate answers.

18. The computer program product of claim **16** wherein the actions further comprise:

identifying a plurality of follow-up postings corresponding to the identified question;

analyzing each of the follow-up postings to identify a conversational move corresponding to each of the follow up postings, wherein at least one of the conversational postings is selected from the group consisting of an answer, a clarification, a rejection, or a different conversational move; and

generating a contribution tree based on the follow-up postings and their identified conversational moves.

19. The computer program product of claim **18** wherein the actions further comprise:

pruning one or more of the follow-up postings from the contribution tree based on a contribution analysis, wherein the pruned follow-up postings have a contribution analysis result selected from the group consisting of an answer leading to a new question, an overly deep follow-up posting, and another pruning criteria.

20. The computer program product of claim **16** wherein the actions further comprise:

wherein at least one of the factors is selected from the group consisting of a quality of the supporting evidence, and a quantity of the supporting evidence;

generating the answer post score based on an identification of a rating within the threaded discussions pertaining to the candidate answer;

generating the post provider score based on an identified expertise level that corresponds to a provider of the candidate answer; and

generating a follow-up score based on one or more follow-up comments from posters that indicate that the candidate answer was correct.

* * * * *